

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2022-214-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2022-214

Anonymous Referee #1

Referee comment on "Towards monitoring CO₂ source-sink distribution over India via inverse modelling: Quantifying the fine-scale spatiotemporal variability of atmospheric CO₂ mole fraction" by Vishnu Thilakan et al., Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2022-214-RC2>, 2022

The study by Thilakan et al. can be understood as a preparatory study for future inverse modelling of CO₂ fluxes over India once an in-situ observation network with sufficient coverage is established. As a starting point, it investigates the spread in CO₂ concentrations simulated by current state-of-the-art global CO₂ data assimilation systems and then analyzes the impact of spatial representation errors in global models on the ability to inversely estimate CO₂ fluxes over India. These representation errors are caused by the inability of coarse global models to resolve small-scale variations in CO₂ due to variations in topography and biospheric and anthropogenic surface fluxes.

The present manuscript is much improved compared to an earlier version and is acceptable with minor revisions. In particular, the motivation of the work is much clearer now, although the individual elements still do not perfectly fit together. The analysis of differences in CO₂ concentrations simulated by different global models is interesting (and in fact alarming), but this part is still only loosely connected to the core of the study, which is the analysis of representation errors. The conclusions section is much more compelling now than in previous versions.

Overall, the study provides a number of interesting analyses that will be valuable for future inversion studies over India and therefore deserves being published. These aspects include

- analysis of the factors contributing to small-scale CO₂ variability
- quantification of corresponding representation errors in global models
- quantification of the impact of these representation errors on inverse estimation of CO₂ fluxes
- presentation of an approach to reduce representation errors in global models, which accounts for the impact of sub-grid scale variability in orography and surface fluxes

I only have a number of minor comments that should be addressed before publication. I trust that other issues in terms of language and grammar will be corrected during the copy-editing phase.

- Line 57: It is true that the predictability of CO₂ concentrations is better for windy situations, but in these situations CO₂ will be dominated by the large background while the sensitivity to regional CO₂ fluxes will be small
- L67: Replace "Further" by "Furthermore"
- L73: Change "over the regions" to "over other regions"
- L94: Replace "availability of radiation" simply by "radiation"
- L126-127: The sentence could be simplified to "The month of July represents a monsoon period when both biospheric and convective activity are strong".
- L129: "In contrast" would probably fit better than "On the other hand"
- Replace "the ways" by "possible ways"
- L206: I was unable to find information on this CAMS product (version 2.2.4). Is this a reanalysis product (like EGG4)? What is the original resolution of this product?
- L207-208: I doubt that the product has a 6-hour temporal and 0.5° spatial resolution. This is probably only the resolution at which the data was obtained, not the resolution of the original product.
- L221: Replace "during the year" by "for the year"
- L223: Delete "for the optimization" because it appears twice in the sentence.
- L238: Simply write "scales not captured" instead of "scales which could not be captured"
- L257-258: Change to "As space-borne instruments measure total columns rather than near-surface concentrations, we extend .."
- L261: Here and at several other places: It would be better to write "errors" rather than "error"
- L266: Change "difference in the sub-grid scale process" to "difference in sub-grid scale processes"
- L272-273: This could be written more elegantly as ".. which describes the systematic component of the representation error and provides important constraints for inversions .."
- L295: Replace "seasonal changes" by "seasonal differences" as you have only analyzed two different months but not the whole year.
- L298: Instead of "associated mesoscale activity." I suggest to write "associated convective activity that is only parameterized in global models".
- L372: Here and later in the same sentence: Replace "grids" by "grid cells"
- L382: Replace "using this approach can be inferred as the lower bound" by "using this approach may be considered as a lower bound"
- L389: Replace "is not sufficient to justify the model's performance" by "does not guarantee a good model performance"
- L404: Here and at a few other places "while" should be replaced by "when" (i.e. "when analyzing")
- L436: The formulation "seasonally varying observations" makes no sense. An observation does not vary seasonally unless e.g. different instruments are used in different seasons.
- L441: Why should the satellite community gap-fill the observations? Standard satellite products do not provide gap-filled data.
- L449: You should also cite Zellweger et al. (2016). <https://amt.copernicus.org/articles/9/4737/2016/>, which shows that the compatibility goal of WMO of 0.1 ppm is achievable with current instruments but is still quite

- challenging. You should also replace "less than 0.1 ppm" by "of the order of 0.1 ppm".
- L460: Replace "over the land and ocean boundary" by "at the boundary between land and ocean"
 - L486: The statement "The estimated column representation error is thus capable of causing significant biases in the satellite inferred CO₂ fluxes over these regions" should be better backed up. How do these errors compare, for example, with the impact of XCO₂ retrieval biases on regional CO₂ fluxes discussed in Chevallier et al. (2007)? <https://doi.org/10.1029/2006JD007375>
 - L570: It should be "the dependence of the representation error on sigma_{bio}" rather than the other way round.
 - L568: It is not quite clear whether the numbers presented here are based on an analysis of the variance or only of the standard deviations. This is important because only variances (and covariances) explained by different factors can be added up to explain the total variance. Standard deviations cannot be added up in this way.
 - L588: It is only the resolution of an atmospheric transport model that is a limitation, not the accuracy of the digital elevation model used to generate the model orography. Differences between different DEMs are typically very small.
 - L597 and following: How does the fact that only 35% of the area of India is covered by the 9 stations affect these results? This should be discussed in this paragraph.
 - L617: How good is the performance in terms of R-square? It would be good to add this information to be consistent with the statement on L613.
 - L624: Again, the accuracy of the DEM is not a limitation.
 - L655: Comparing Figures 4 and 5 or looking at Figure 6, I don't see an order of magnitude difference in representation errors between surface and column CO₂. The representation errors in column CO₂ are in fact surprisingly large.